

sulphate; and EDTA is used as an aminotetracetic acid. After hybridization, the membrane upon which the DNA is transferred is washed at 2 x SSC at room temperature and then at 0.1 x SSC/0.1 x SDS at temperatures up to 68°C.

There are other conditions, reagents, and so forth which can be used, and would result in a similar degree of stringency. The skilled artisan will be familiar with such conditions, and thus they are not given here. It will be understood, however, that the skilled artisan will be able to manipulate the conditions in a manner to permit the clear identification of homologs and alleles of LEDGF nucleic acids of the invention. The skilled artisan also is familiar with the methodology for screening cells and libraries for expression of such molecules which then are routinely isolated, followed by isolation of the pertinent nucleic acid molecule and sequencing.

In general homologs and alleles typically will share at least 40% nucleotide identity and/or at least 50% amino acid identity to SEQ ID NO:1 and SEQ ID NO:2, respectively, in some instances will share at least 50% nucleotide identity and/or at least 65% amino acid identity and in still other instances will share at least 60% nucleotide identity and/or at least 75% amino acid identity. The homology can be calculated using various, publicly available software tools developed by NCBI (Bethesda, Maryland) that can be obtained through the internet. Exemplary tools include the BLAST system. Pairwise and ClustalW alignments (BLOSUM30 matrix setting) as well as Kyte-Doolittle hydropathic analysis can be obtained using the MacVetor sequence analysis software (Oxford Molecular Group). Watson-Crick complements of the foregoing nucleic acids also are embraced by the invention.

In screening for LEDGF related genes, such as homologs and alleles of LEDGF, a Southern blot may be performed using the foregoing conditions, together with a radioactive probe. After washing the membrane to which the DNA is finally transferred, the membrane can be placed against X-ray film or a phosphorimager plate to detect the radioactive signal.

Given that the expression of the LEDGF gene is abundant in certain human tissues, and given the teachings herein of a full-length human LEDGF cDNA clone, other mammalian sequences such as the mouse cDNA clone corresponding to the human LEDGF gene can be isolated from a cDNA library prepared from one or more of the tissues in which LEDGF expression is abundant, using standard colony hybridization techniques.

The invention also includes degenerate nucleic acids which include alternative codons

isolated from a cDNA library prepared from one or more of the tissues in which LEDGF expression is abundant, using standard colony hybridization techniques.

The invention also includes degenerate nucleic acids which include alternative codons